

SEQUENCE LISTING

<110> Università degli Studi di Roma "La Sapienza"
Consiglio Nazionale delle Ricerche

<120> Purification, cloning and biochemical characterization of XendoU,
endoribonucleasic activity involved in *Xenopus laevis* small
nuclear RNA splicing independent biosynthesis

<130> PCT25390

<150> IT RM2002A000365
<151> 2002-07-08

<160> 9

<170> PatentIn version 3.2

<210> 1
<211> 1265
<212> DNA
<213> *Xenopus laevis*

<220>
<221> misc_feature
<222> (1)..(38)
<223> 5' untranslated DNA region

<220>
<221> exon
<222> (39)..(915)

<220>
<221> terminator
<222> (916)..(918)

<220>
<221> misc_feature
<222> (916)..(1265)
<223> 3' untranslated DNA region

<400> 1
attggggaac tgggagcaga gagtgacggg caggagcc atg gcg agt aac agg ggg 56
Met Ala Ser Asn Arg Gly
1 5

cag ctg aac cat gaa ctc tcc aag ctg ttt aat gag ctg tgg gac gca 104
Gln Leu Asn His Glu Leu Ser Lys Leu Phe Asn Glu Leu Trp Asp Ala
10 15 20

gat cag aac cgg atg aag tcc ggg aag gat tat cgg atc tcc ttg cag 152
Asp Gln Asn Arg Met Lys Ser Gly Lys Asp Tyr Arg Ile Ser Leu Gln
25 30 35

ggt aaa gca ggg tac gta ccc gcc ggt tcc aac cag gcc agg gac agc 200
Gly Lys Ala Gly Tyr Val Pro Ala Gly Ser Asn Gln Ala Arg Asp Ser

| 40 | 45 | 50 | |
|---|----|----|-----|
| gcc tcg ttc ccg ctc ttc cag ttc gtc gat gag gag aag ctg aag agc Ala Ser Phe Pro Leu Phe Gln Phe Val Asp Glu Glu Lys Leu Lys Ser 55 60 65 70 | | | 248 |
| agg aag acg ttt gca acc ttc att tcc ctg ctg gac aat tat gag atg Arg Lys Thr Phe Ala Thr Phe Ile Ser Leu Leu Asp Asn Tyr Glu Met 75 80 85 | | | 296 |
| gac acg ggg gtg gcc gag gtt gtg act ccg gag gaa atc gct gaa aac Asp Thr Gly Val Ala Glu Val Val Thr Pro Glu Glu Ile Ala Glu Asn 90 95 100 | | | 344 |
| aac aac ttc ctg gac gcc att ctg gaa acc aaa gtg atg aag atg gca Asn Asn Phe Leu Asp Ala Ile Leu Glu Thr Lys Val Met Lys Met Ala 105 110 115 | | | 392 |
| cat gac tac ctg gtg agg aag aac caa gcc aaa ccc acc cgg aat gac His Asp Tyr Leu Val Arg Lys Asn Gln Ala Lys Pro Thr Arg Asn Asp 120 125 130 | | | 440 |
| ttc aag gtc caa ctg tac aac atc tgg ttc cag ctg tac tca cgg gcc Phe Lys Val Gln Leu Tyr Asn Ile Trp Phe Gln Leu Tyr Ser Arg Ala 135 140 145 150 | | | 488 |
| cca ggg agc aga ccc gat tcg tgc ggc ttt gag cac gtg ttt gtg gga Pro Gly Ser Arg Pro Asp Ser Cys Gly Phe Glu His Val Phe Val Gly 155 160 165 | | | 536 |
| gaa tcg aag cga ggg cag gag atg atg ggg ctt cac aac tgg gtc cag Glu Ser Lys Arg Gly Gln Glu Met Met Gly Leu His Asn Trp Val Gln 170 175 180 | | | 584 |
| ttt tac ctt cag gag aag agg aag aac atc gac tat aaa gga tac gtg Phe Tyr Leu Gln Glu Lys Arg Lys Asn Ile Asp Tyr Lys Gly Tyr Val 185 190 195 | | | 632 |
| gct cgg cag aac aag agt cgg ccg gat gaa gat gat cag gtg ttg aac Ala Arg Gln Asn Lys Ser Arg Pro Asp Glu Asp Asp Gln Val Leu Asn 200 205 210 | | | 680 |
| ctg cag ttc aat tgg aag gag atg gtg aaa ccc gtc ggc agc agc ttc Leu Gln Phe Asn Trp Lys Glu Met Val Lys Pro Val Gly Ser Ser Phe 215 220 225 230 | | | 728 |
| att ggc gtc agc ccg gaa ttc gaa ttc gcc ctt tac acc atc gtc ttc Ile Gly Val Ser Pro Glu Phe Glu Phe Ala Leu Tyr Thr Ile Val Phe 235 240 245 | | | 776 |
| ctc gcg tct cag gag aag atg agc cga gaa gtc gtt cgg ctg gaa gaa Leu Ala Ser Gln Glu Lys Met Ser Arg Glu Val Val Arg Leu Glu Glu 250 255 260 | | | 824 |
| tac gaa ctg cag atc gtc gtc aat cgc cac ggc cgt tat ata ggg acc Tyr Glu Leu Gln Ile Val Val Asn Arg His Gly Arg Tyr Ile Gly Thr 265 270 275 | | | 872 |

gcc tac ccc gtc ctc ctg agc acc aat aac ccg gat ctg tac t 915
 Ala Tyr Pro Val Leu Leu Ser Thr Asn Asn Pro Asp Leu Tyr
 280 285 290

gagggggcgg ggctagagat cacagccggt tcccacgggt tgggtgcatt tactaacaaa 975

actgcaccaa tgcaacaatg caagcagata atgggggcag gtccatatcc ctctgctttc 1035

cctagcgtgt gtggggcaca ttaaccctat aactgtcact cactgcacca gaccattat 1095

ttaacccac aaggacatc aagccagtgc cttgttatga gagagcgcag ccggggcttc 1155

tctactgtga aacttctgta ttgtatagag tttacttgggt ttcttctcc agacaatttc 1215

actttttttt tgctttgcct ttaaccatta aaagtccatg acatttctgt 1265

<210> 2

<211> 292

<212> PRT

<213> *Xenopus laevis*

<400> 2

Met Ala Ser Asn Arg Gly Gln Leu Asn His Glu Leu Ser Lys Leu Phe
 1 5 10 15

Asn Glu Leu Trp Asp Ala Asp Gln Asn Arg Met Lys Ser Gly Lys Asp
 20 25 30

Tyr Arg Ile Ser Leu Gln Gly Lys Ala Gly Tyr Val Pro Ala Gly Ser
 35 40 45

Asn Gln Ala Arg Asp Ser Ala Ser Phe Pro Leu Phe Gln Phe Val Asp
 50 55 60

Glu Glu Lys Leu Lys Ser Arg Lys Thr Phe Ala Thr Phe Ile Ser Leu
 65 70 75 80

Leu Asp Asn Tyr Glu Met Asp Thr Gly Val Ala Glu Val Val Thr Pro
 85 90 95

Glu Glu Ile Ala Glu Asn Asn Asn Phe Leu Asp Ala Ile Leu Glu Thr
 100 105 110

Lys Val Met Lys Met Ala His Asp Tyr Leu Val Arg Lys Asn Gln Ala
 115 120 125

Lys Pro Thr Arg Asn Asp Phe Lys Val Gln Leu Tyr Asn Ile Trp Phe
 130 135 140

Gln Leu Tyr Ser Arg Ala Pro Gly Ser Arg Pro Asp Ser Cys Gly Phe
 145 150 155 160

Arg His Val Phe Val Gly Glu Ser Lys Arg Gly Gln Glu Met Met Gly
 165 170 175

Leu His Asn Trp Val Gln Phe Tyr Leu Gln Glu Lys Arg Lys Asn Ile
 180 185 190

Asp Tyr Lys Gly Tyr Val Ala Arg Gln Asn Lys Ser Arg Pro Asp Glu
 195 200 205

Asp Asp Gln Val Leu Asn Leu Gln Phe Asn Trp Lys Glu Met Val Lys
 210 215 220

Pro Val Gly Ser Ser Phe Ile Gly Val Ser Pro Glu Phe Glu Phe Ala
 225 230 235 240

Leu Tyr Thr Ile Val Phe Leu Ala Ser Gln Glu Lys Met Ser Arg Glu
 245 250 255

Val Val Arg Leu Glu Glu Tyr Glu Leu Gln Ile Val Val Asn Arg His
 260 265 270

Gly Arg Tyr Ile Gly Thr Ala Tyr Pro Val Leu Leu Ser Thr Asn Asn
 275 280 285

Pro Asp Leu Tyr
 290

<210> 3
 <211> 20
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligoribonucleotide

<220>
 <221> misc_RNA
 <222> (1)..(20)
 <223> Synthetic oligoribonucleotide which includes U16 upstream

cleavage site

<400> 3
ggaaacguau ccuuugggag

20

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Shyntetic oligonucleotide

<220>
<221> misc_RNA
<222> (1)..(20)
<223> Mutant of SEQ ID 3

<220>
<221> variation
<222> (20)..(20)
<223> "U" is replaced by "T"

<400> 4
ggaaacguau ccuugggagt

20

<210> 5
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic oligoribonucleotide

<220>
<221> misc_RNA
<222> (1)..(20)
<223> Mutant of SEQ ID 3

<220>
<221> variation
<222> (14)..(14)
<223> "U" is replaced by "C"

<400> 5
ggaaacguau ccucugggag

20

<210> 6
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic oligoribonucleotide

<220>
<221> misc_RNA
<222> (1)..(20)
<223> Mutant of SEQ ID 3

<220>
<221> variation
<222> (14)..(14)
<223> "U" is replaced by "G"

<400> 6
ggaaacguau ccugugggag

20

<210> 7
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<220>
<221> misc_feature
<222> (1)..(27)
<223> Primer

<400> 7
aagcttcttc atggcggctc ggccaat

27

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> MAHs: synthetic degenerate oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer element

<220>
<221> modified_base
<222> (6)..(6)
<223> "n" : I

<220>
<221> modified_base

<222> (18)..(18)
<223> "n" : I

<400> 8
atggcncayg aytayytngt

20

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> IGTa: synthetic degenerate oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer element

<220>
<221> modified_base
<222> (3)..(3)
<223> "n" : I

<220>
<221> modified_base
<222> (9)..(9)
<223> "n" : I

<220>
<221> modified_base
<222> (12)..(12)
<223> "n": I

<220>
<221> modified_base
<222> (15)..(15)
<223> "n" : I

<220>
<221> modified_base
<222> (18)..(18)
<223> "n" : I

<400> 9
acnggrtang cngtnccnat

20